

RADOS - Bug #38195

osd-backfill-space.sh exposes rocksdb hang

02/05/2019 07:01 PM - David Zafman

Status: New	% Done: 0%
Priority: Normal	Spent time: 0.00 hour
Assignee:	
Category:	
Target version:	
Source:	Affected Versions:
Tags:	ceph-qa-suite:
Backport:	Component(RADOS):
Regression: No	Pull request ID:
Severity: 3 - minor	Crash signature (v1):
Reviewed:	Crash signature (v2):
Description	
After increasing the timeout for backfills finish for http://tracker.ceph.com/issues/38027 we see this kill_daemons() error with ceph-mon.	
dzafman-2019-02-04_11:24:54-rados-wip-zafman-testing2-distro-basic-smithi/3549933	
Before mon shutdowns it appears to hang in rocksdb before SIGKILL sent by teardown. This run exposes the kill_daemons() execution.	
teuthology.log	
<pre>2019-02-04T20:24:27.598 INFO:tasks.workunit.client.0.smithi194.stderr:35707: /home/ubuntu/ceph-test/clone.c35707: //home/ubuntu/ceph-test/clone.client.0/qa/standalone/ceph-helpers.sh:255: kill_daemon: cat td/osd-backfill-space/mon.a.pid 2019-02-04T20:24:27.598 INFO:tasks.workunit.client.0.smithi194.stderr:35707: /home/ubuntu/ceph-test/clone.client.0/qa/standalone/ceph-helpers.sh:255: kill_daemon: local pid=126389 2019-02-04T20:24:27.598 INFO:tasks.workunit.client.0.smithi194.stderr:35707: /home/ubuntu/ceph-test/clone.client.0/qa/standalone/ceph-helpers.sh:256: kill_daemon: local send_signal=TERM 2019-02-04T20:24:27.598 INFO:tasks.workunit.client.0.smithi194.stderr:35707: /home/ubuntu/ceph-test/clone.client.0/qa/standalone/ceph-helpers.sh:257: kill_daemon: local 'delays=0.1 0.2 1 1 1 2 3 5 5 5 10 10 20 60 60 60 120' 2019-02-04T20:24:27.598 INFO:tasks.workunit.client.0.smithi194.stderr:35707: /home/ubuntu/ceph-test/clone.client.0/qa/standalone/ceph-helpers.sh:258: kill_daemon: local exit_code=1 2019-02-04T20:24:27.598 INFO:tasks.workunit.client.0.smithi194.stderr:35707: /home/ubuntu/ceph-test/clone.client.0/qa/standalone/ceph-helpers.sh:259: kill_daemon: for try in '\$delays' 2019-02-04T20:30:25.590 INFO:tasks.workunit.client.0.smithi194.stderr:35707: /home/ubuntu/ceph-test/clone.client.0/qa/standalone/ceph-helpers.sh:259: kill_daemon: for try in '\$delays' 2019-02-04T20:30:25.591 INFO:tasks.workunit.client.0.smithi194.stderr:35707: /home/ubuntu/ceph-test/clone.client.0/qa/standalone/ceph-helpers.sh:260: kill_daemon: kill -0 126389 2019-02-04T20:30:25.591 INFO:tasks.workunit.client.0.smithi194.stderr:35707: /home/ubuntu/ceph-test/clone.client.0/qa/standalone/ceph-helpers.sh:261: kill_daemon: exit_code=1 2019-02-04T20:30:25.591 INFO:tasks.workunit.client.0.smithi194.stderr:35707: /home/ubuntu/ceph-test/clone.client.0/qa/standalone/ceph-helpers.sh:266: kill_daemon: send_signal=0 2019-02-04T20:30:25.591 INFO:tasks.workunit.client.0.smithi194.stderr:35707: /home/ubuntu/ceph-test/clone.client.0/qa/standalone/ceph-helpers.sh:267: kill_daemon: sleep 120 2019-02-04T20:30:25.591 INFO:tasks.workunit.client.0.smithi194.stderr:35707: /home/ubuntu/ceph-test/clone.client.0/qa/standalone/ceph-helpers.sh:269: kill_daemon: return 1 2019-02-04T20:30:25.594 INFO:tasks.workunit.client.0.smithi194.stderr:/home/ubuntu/ceph-test/clone.client.0/qa/standalone/ceph-helpers.sh:1866: run_in_background: return 1</pre>	

mon.a.log

```
...
2019-02-04 20:24:25.296 7f1b39312700 4 rocksdb: (Original Log Time 2019/02/04-20:24:25.257518) [/
build/ceph-14.0.1-3131-ge01bb11/src/rocksdb/db/memtable_list.cc:377] [default] Level-0 commit tabl
e #19 started
2019-02-04 20:24:25.296 7f1b39312700 4 rocksdb: (Original Log Time 2019/02/04-20:24:25.299890) [/
build/ceph-14.0.1-3131-ge01bb11/src/rocksdb/db/memtable_list.cc:424] [default] Level-0 commit tabl
e #19: memtable #1 done
2019-02-04 20:24:25.296 7f1b39312700 4 rocksdb: (Original Log Time 2019/02/04-20:24:25.299929) EV
ENT_LOG_v1 {"time_micros": 1549311865299904, "job": 9, "event": "flush_finished", "output_compress
ion": "NoCompression", "lsm_state": [1, 0, 0, 0, 0, 0, 1], "immutable_memtables": 0}
2019-02-04 20:24:25.296 7f1b39312700 4 rocksdb: (Original Log Time 2019/02/04-20:24:25.299950) [/
build/ceph-14.0.1-3131-ge01bb11/src/rocksdb/db/db_impl_compaction_flush.cc:177] [default] Level su
mmmary: base level 6 max bytes base 268435456 files[1 0 0 0 0 0 1] max score 0.25

2019-02-04 20:24:25.296 7f1b39312700 4 rocksdb: [/build/ceph-14.0.1-3131-ge01bb11/src/rocksdb/db/
db_impl_files.cc:347] [JOB 9] Try to delete WAL files size 9979119, prev total WAL file size 10025
925, number of live WAL files 2.

2019-02-04 20:24:25.296 7f1b30300700 4 rocksdb: [/build/ceph-14.0.1-3131-ge01bb11/src/rocksdb/db/
db_impl_compaction_flush.cc:1060] [default] Manual compaction starting
2019-02-04 20:24:25.296 7f1b39b13700 4 rocksdb: (Original Log Time 2019/02/04-20:24:25.300155) [/
build/ceph-14.0.1-3131-ge01bb11/src/rocksdb/db/db_impl_compaction_flush.cc:1847] [default] Manual
compaction from level-0 to level-6 from 'paxos .. 'paxos; will stop at (end)

2019-02-04 20:24:25.296 7f1b39b13700 4 rocksdb: [/build/ceph-14.0.1-3131-ge01bb11/src/rocksdb/db/
compaction_job.cc:1636] [default] [JOB 10] Compacting 1@0 + 1@6 files to L6, score -1.00
2019-02-04 20:24:25.296 7f1b39b13700 4 rocksdb: [/build/ceph-14.0.1-3131-ge01bb11/src/rocksdb/db/
compaction_job.cc:1640] [default] Compaction start summary: Base version 9 Base level 0, inputs: [
19(9580KB)], [17(18MB)]

2019-02-04 20:24:25.296 7f1b39b13700 4 rocksdb: EVENT_LOG_v1 {"time_micros": 1549311865300218, "j
ob": 10, "event": "compaction_started", "compaction_reason": "ManualCompaction", "files_L0": [19],
"files_L6": [17], "score": -1, "input_data_size": 29158732}
2019-02-04 20:24:25.524 7f1b39b13700 4 rocksdb: [/build/ceph-14.0.1-3131-ge01bb11/src/rocksdb/db/
compaction_job.cc:1324] [default] [JOB 10] Generated table #20: 2415 keys, 25628909 bytes
2019-02-04 20:24:25.524 7f1b39b13700 4 rocksdb: EVENT_LOG_v1 {"time_micros": 1549311865525468, "c
f_name": "default", "job": 10, "event": "table_file_creation", "file_number": 20, "file_size": 256
28909, "table_properties": {"data_size": 25510081, "index_size": 52487, "filter_size": 65628, "raw
_key_size": 44917, "raw_average_key_size": 18, "raw_value_size": 25439814, "raw_average_value_size
": 10534, "num_data_blocks": 1603, "num_entries": 2415, "filter_policy_name": "rocksdb.BuiltinBloo
mFilter", "kDeletedKeys": "0", "kMergeOperands": "0"}}
2019-02-04 20:24:25.548 7f1b39b13700 4 rocksdb: [/build/ceph-14.0.1-3131-ge01bb11/src/rocksdb/db/
compaction_job.cc:1390] [default] [JOB 10] Compacted 1@0 + 1@6 files to L6 => 25628909 bytes
2019-02-04 20:24:25.548 7f1b39b13700 3 rocksdb: [/build/ceph-14.0.1-3131-ge01bb11/src/rocksdb/db/
version_set.cc:2578] More existing levels in DB than needed. max_bytes_for_level_multiplier may no
t be guaranteed.
2019-02-04 20:24:25.616 7f1b39b13700 4 rocksdb: (Original Log Time 2019/02/04-20:24:25.617853) [/
build/ceph-14.0.1-3131-ge01bb11/src/rocksdb/db/compaction_job.cc:752] [default] compacted to: base
level 6 max bytes base 268435456 files[0 0 0 0 0 0 1] max score 0.00, MB/sec: 129.4 rd, 113.8 wr,
level 6, files in(1, 1) out(1) MB in(9.4, 18.5) out(24.4), read-write-amplify(5.6) write-amplify(
2.6) OK, records in: 2929, records dropped: 514 output_compression: NoCompression

2019-02-04 20:24:25.616 7f1b39b13700 4 rocksdb: (Original Log Time 2019/02/04-20:24:25.617896) EV
ENT_LOG_v1 {"time_micros": 1549311865617881, "job": 10, "event": "compaction_finished", "compactio
n_time_micros": 225308, "output_level": 6, "num_output_files": 1, "total_output_size": 25628909, "
num_input_records": 2929, "num_output_records": 2415, "num_subcompactions": 1, "output_compression
": "NoCompression", "num_single_delete_mismatches": 0, "num_single_delete_fallthrough": 0, "lsm_st
ate": [0, 0, 0, 0, 0, 1]}
2019-02-04 20:24:25.616 7f1b39b13700 4 rocksdb: EVENT_LOG_v1 {"time_micros": 1549311865619390, "j
ob": 10, "event": "table_file_deletion", "file_number": 19}
2019-02-04 20:24:25.620 7f1b39b13700 4 rocksdb: EVENT_LOG_v1 {"time_micros": 1549311865621737, "j
ob": 10, "event": "table_file_deletion", "file_number": 17}
```

```
2019-02-04 20:24:25.620 7f1b30300700 4 rocksdb: [/build/ceph-14.0.1-3131-ge01bb11/src/rocksdb/db/db_impl_compaction_flush.cc:1060] [default] Manual compaction starting
2019-02-04 20:24:25.620 7f1b30300700 4 rocksdb: [/build/ceph-14.0.1-3131-ge01bb11/src/rocksdb/db/db_impl_compaction_flush.cc:1060] [default] Manual compaction starting
2019-02-04 20:24:25.620 7f1b30300700 4 rocksdb: [/build/ceph-14.0.1-3131-ge01bb11/src/rocksdb/db/db_impl_compaction_flush.cc:1060] [default] Manual compaction starting
2019-02-04 20:24:25.620 7f1b30300700 4 rocksdb: [/build/ceph-14.0.1-3131-ge01bb11/src/rocksdb/db/db_impl_compaction_flush.cc:1060] [default] Manual compaction starting
2019-02-04 20:24:25.620 7f1b30300700 4 rocksdb: [/build/ceph-14.0.1-3131-ge01bb11/src/rocksdb/db/db_impl_compaction_flush.cc:1060] [default] Manual compaction starting
<END OF LOG>
```

Related issues:

Related to RADOS - Bug #38027: osd/osd-backfill-space.sh fails

Resolved

01/23/2019

History

#1 - 02/05/2019 07:01 PM - David Zafman

- Related to Bug #38027: osd/osd-backfill-space.sh fails added

#2 - 02/05/2019 07:03 PM - David Zafman

- Status changed from New to 12

- Priority changed from Normal to Urgent

After adding code to send a SEGV on kill_daemon timeout I got the following stack traces.

```
(gdb) thread apply all bt
```

```
Thread 13 (Thread 0x7f04e291b700 (LWP 151409)):
#0 0x00007f04e7b1420d in poll () from /lib64/libc.so.6
#1 0x000055c9ccc24599 in poll (__timeout=-1, __nfds=4, __fds=0x7f04e29183e0) at /usr/include/bits/poll2.h:41
#2 SignalHandler::entry (this=0x55c9cf9d6dc0) at /usr/src/debug/ceph-14.0.1-3134-gdfall176/src/global/signal_handler.cc:459
#3 0x00007f04e8c58dd5 in start_thread () from /lib64/libpthread.so.0
#4 0x00007f04e7b1eead in clone () from /lib64/libc.so.6
```

```
Thread 12 (Thread 0x7f04e1118700 (LWP 151412)):
#0 0x00007f04e8c5c965 in pthread_cond_wait@@GLIBC_2.3.2 () from /lib64/libpthread.so.0
#1 0x000055c9cce25d9d in rocksdb::port::CondVar::Wait (this=this@entry=0x55c9d081b350) at /usr/src/debug/ceph-14.0.1-3134-gdfall176/src/rocksdb/port/port_posix.cc:91
#2 0x000055c9ccd80775 in rocksdb::InstrumentedCondVar::WaitInternal (this=this@entry=0x55c9d081b350) at /usr/src/debug/ceph-14.0.1-3134-gdfall176/src/rocksdb/monitoring/instrumented_mutex.cc:48
#3 0x000055c9ccd80852 in rocksdb::InstrumentedCondVar::Wait (this=this@entry=0x55c9d081b350) at /usr/src/debug/ceph-14.0.1-3134-gdfall176/src/rocksdb/monitoring/instrumented_mutex.cc:41
#4 0x000055c9cce05110 in rocksdb::DeleteScheduler::BackgroundEmptyTrash (this=0x55c9d081b260) at /usr/src/debug/ceph-14.0.1-3134-gdfall176/src/rocksdb/util/delete_scheduler.cc:193
#5 0x00007f04ec7cec5f in execute_native_thread_routine () from /usr/lib64/ceph/libceph-common.so.1
#6 0x00007f04e8c58dd5 in start_thread () from /lib64/libpthread.so.0
#7 0x00007f04e7b1eead in clone () from /lib64/libc.so.6
```

```
Thread 11 (Thread 0x7f04e0917700 (LWP 151432)):
#0 0x00007f04e7ae5e2d in nanosleep () from /lib64/libc.so.6
#1 0x00007f04e7b16704 in usleep () from /lib64/libc.so.6
#2 0x00007f04ec323b6b in OpHistoryServiceThread::entry (this=0x55c9d09f3190) at /usr/src/debug/ceph-14.0.1-3134-gdfall176/src/common/TrackedOp.cc:44
#3 0x00007f04e8c58dd5 in start_thread () from /lib64/libpthread.so.0
#4 0x00007f04e7b1eead in clone () from /lib64/libc.so.6
```

```
Thread 10 (Thread 0x7f04d8907700 (LWP 212750)):
#0 0x00007f04e8c5c965 in pthread_cond_wait@@GLIBC_2.3.2 () from /lib64/libpthread.so.0
#1 0x000055c9ccaf01b7 in Wait (mutex=..., this=0x55c9cfaf14e50) at /usr/src/debug/ceph-14.0.1-3134-gdfall176/src/common/Cond.h:49
#2 RocksDBStore::compact_thread_entry (this=0x55c9cfaf14c80) at /usr/src/debug/ceph-14.0.1-3134-gdfall176/src/kv/RocksDBStore.cc:1223
#3 0x000055c9ccae3b9d in RocksDBStore::CompactThread::entry (this=<optimized out>) at /usr/src/debug/ceph-14.0.1-3134-gdfall176/src/kv/RocksDBStore.h:107
#4 0x00007f04e8c58dd5 in start_thread () from /lib64/libpthread.so.0
#5 0x00007f04e7b1eead in clone () from /lib64/libc.so.6
```

```

Thread 9 (Thread 0x7f04e41a6700 (LWP 151407)):
#0 0x00007f04e8c5cd12 in pthread_cond_timedwait@@GLIBC_2.3.2 () from /lib64/libpthread.so.0
#1 0x00007f04ec367f89 in __gthread_cond_timedwait (__abs_timeout=0x7f04e41a3640, __mutex=<optimized out>, __cond=0x55c9cf9e9f38) at /opt/rh/devtoolset-7/root/usr/include/c++/7/x86_64-redhat-linux/bits/gthr-default.h:871
#2 __wait_until_impl<std::chrono::duration<unsigned long, std::ratio<1, 1000000000> > > (&__atime=..., __lock=..., this=0x55c9cf9e9f38) at /opt/rh/devtoolset-7/root/usr/include/c++/7/condition_variable:166
#3 wait_until<std::chrono::duration<unsigned long, std::ratio<1, 1000000000> > > (&__atime=..., __lock=..., this=0x55c9cf9e9f38) at /opt/rh/devtoolset-7/root/usr/include/c++/7/condition_variable:106
#4 wait_for<unsigned long, std::ratio<1, 1000000000> > (&__rtime=<synthetic pointer>, __lock=..., this=0x55c9cf9e9f38) at /opt/rh/devtoolset-7/root/usr/include/c++/7/condition_variable:138
#5 CephContextServiceThread::entry (this=0x55c9cf9e9ef0) at /usr/src/debug/ceph-14.0.1-3134-gdfall176/src/common/ceph_context.cc:197
#6 0x00007f04e8c58dd5 in start_thread () from /lib64/libpthread.so.0
#7 0x00007f04e7b1eead in clone () from /lib64/libc.so.6

Thread 8 (Thread 0x7f04e39a5700 (LWP 151408)):
#0 0x00007f04e7b1420d in poll () from /lib64/libc.so.6
#1 0x00007f04ec34697b in poll (__timeout=-1, __nfds=2, __fds=0x7f04e39a2620) at /usr/include/bits/poll2.h:46
#2 AdminSocket::entry (this=0x55c9cf9eaf70) at /usr/src/debug/ceph-14.0.1-3134-gdfall176/src/common/admin_socket.cc:228
#3 0x00007f04ec7cec5f in execute_native_thread_routine () from /usr/lib64/ceph/libceph-common.so.1
#4 0x00007f04e8c58dd5 in start_thread () from /lib64/libpthread.so.0
#5 0x00007f04e7b1eead in clone () from /lib64/libc.so.6

Thread 7 (Thread 0x7f04e1919700 (LWP 151411)):
#0 0x00007f04e8c5c965 in pthread_cond_wait@@GLIBC_2.3.2 () from /lib64/libpthread.so.0
#1 0x00007f04e83b782c in std::condition_variable::wait(std::unique_lock<std::mutex&>) () from /lib64/libstdc++.so.6
#2 0x000055c9ccea649 in rocksdb::ThreadPoolImpl::Impl::BGThread (this=this@entry=0x55c9cf9f7420, thread_id=thread_id@entry=0) at /usr/src/debug/ceph-14.0.1-3134-gdfall176/src/rocksdb/util/threadpool_imp.cc:196
#3 0x000055c9ccea97f in rocksdb::ThreadPoolImpl::Impl::BGThreadWrapper (arg=0x55c9cf9d05d0) at /usr/src/debug/ceph-14.0.1-3134-gdfall176/src/rocksdb/util/threadpool_imp.cc:303
#4 0x00007f04ec7cec5f in execute_native_thread_routine () from /usr/lib64/ceph/libceph-common.so.1
#5 0x00007f04e8c58dd5 in start_thread () from /lib64/libpthread.so.0
#6 0x00007f04e7b1eead in clone () from /lib64/libc.so.6

Thread 6 (Thread 0x7f04db10c700 (LWP 151431)):
#0 0x00007f04e7b1f483 in epoll_wait () from /lib64/libc.so.6
#1 0x00007f04ec5881c9 in EpollDriver::event_wait (this=0x55c9d08e8cc0, fired_events=std::vector of length 0, capacity 0, tvp=<optimized out>) at /usr/src/debug/ceph-14.0.1-3134-gdfall176/src/msg/async/EventPoll.cc:122
#2 0x00007f04ec57f35d in EventCenter::process_events (this=this@entry=0x55c9cfa31880, timeout_microseconds=<optimized out>, timeout_microseconds@entry=30000000, working_dur=working_dur@entry=0x7f04db1095e0) at /usr/src/debug/ceph-14.0.1-3134-gdfall176/src/msg/async/Event.cc:401
#3 0x00007f04ec584815 in operator() (__closure=0x55c9d08e8548) at /usr/src/debug/ceph-14.0.1-3134-gdfall176/src/msg/async/Stack.cc:53
#4 std::Function_handler<void(), NetworkStack::add_thread(unsigned int)::<lambda()> >::_M_invoke(const std::_Any_data &) (__functor=...) at /opt/rh/devtoolset-7/root/usr/include/c++/7/bits/std_function.h:316
#5 0x00007f04ec7cec5f in execute_native_thread_routine () from /usr/lib64/ceph/libceph-common.so.1
#6 0x00007f04e8c58dd5 in start_thread () from /lib64/libpthread.so.0
#7 0x00007f04e7b1eead in clone () from /lib64/libc.so.6

Thread 5 (Thread 0x7f04e211a700 (LWP 151410)):
#0 0x00007f04e8c5c965 in pthread_cond_wait@@GLIBC_2.3.2 () from /lib64/libpthread.so.0
#1 0x00007f04e83b782c in std::condition_variable::wait(std::unique_lock<std::mutex&>) () from /lib64/libstdc++.so.6
#2 0x000055c9ccea649 in rocksdb::ThreadPoolImpl::Impl::BGThread (this=this@entry=0x55c9cf9f7340, thread_id=thread_id@entry=0) at /usr/src/debug/ceph-14.0.1-3134-gdfall176/src/rocksdb/util/threadpool_imp.cc:196
#3 0x000055c9ccea97f in rocksdb::ThreadPoolImpl::Impl::BGThreadWrapper (arg=0x55c9cf9d05c0) at /usr/src/debug/ceph-14.0.1-3134-gdfall176/src/rocksdb/util/threadpool_imp.cc:303
#4 0x00007f04ec7cec5f in execute_native_thread_routine () from /usr/lib64/ceph/libceph-common.so.1
#5 0x00007f04e8c58dd5 in start_thread () from /lib64/libpthread.so.0
---Type <return> to continue, or q <return> to quit---
#6 0x00007f04e7b1eead in clone () from /lib64/libc.so.6

Thread 4 (Thread 0x7f04e5123700 (LWP 151406)):
#0 0x00007f04e8c5c965 in pthread_cond_wait@@GLIBC_2.3.2 () from /lib64/libpthread.so.0
#1 0x00007f04e83b782c in std::condition_variable::wait(std::unique_lock<std::mutex&>) () from /lib64/libstdc++.so.6
#2 0x00007f04ec5b8adf in ceph::logging::Log::entry (this=0x55c9cf9ee960) at /usr/src/debug/ceph-14.0.1-3134-gdfall176/src/log/Log.cc:418
#3 0x00007f04e8c58dd5 in start_thread () from /lib64/libpthread.so.0
#4 0x00007f04e7b1eead in clone () from /lib64/libc.so.6

Thread 3 (Thread 0x7f04da10a700 (LWP 151429)):
#0 0x00007f04e7b1f483 in epoll_wait () from /lib64/libc.so.6

```

```
#1 0x00007f04ec5881c9 in EpollDriver::event_wait (this=0x55c9d08e8db0, fired_events=std::vector of length 0, capacity 0, tvp=<optimized out>) at /usr/src/debug/ceph-14.0.1-3134-gdfall176/src/msg/async/EventEpoll.cc:122
#2 0x00007f04ec57f35d in EventCenter::process_events (this=this@entry=0x55c9cfa31080, timeout_microseconds=<optimized out>, timeout_microseconds@entry=30000000, working_dur=working_dur@entry=0x7f04da1075e0) at /usr/src/debug/ceph-14.0.1-3134-gdfall176/src/msg/async/Event.cc:401
#3 0x00007f04ec584815 in operator() (__closure=0x55c9d08e8f68) at /usr/src/debug/ceph-14.0.1-3134-gdfall176/src/msg/async/Stack.cc:53
#4 std::_Function_handler<void(), NetworkStack::add_thread(unsigned int)::<lambda()>>::_M_invoke(const std::_Any_data &) (__functor=...) at /opt/rh/devtoolset-7/root/usr/include/c++/7/bits/std_function.h:316
#5 0x00007f04ec7cec5f in execute_native_thread_routine () from /usr/lib64/ceph/libceph-common.so.1
#6 0x00007f04e8c58dd5 in start_thread () from /lib64/libpthread.so.0
#7 0x00007f04e7b1eead in clone () from /lib64/libc.so.6
```

Thread 2 (Thread 0x7f04da90b700 (LWP 151430)):

```
#0 0x00007f04e7b1f483 in epoll_wait () from /lib64/libc.so.6
#1 0x00007f04ec5881c9 in EpollDriver::event_wait (this=0x55c9d08e8d80, fired_events=std::vector of length 0, capacity 0, tvp=<optimized out>) at /usr/src/debug/ceph-14.0.1-3134-gdfall176/src/msg/async/EventEpoll.cc:122
#2 0x00007f04ec57f35d in EventCenter::process_events (this=this@entry=0x55c9cfa31480, timeout_microseconds=<optimized out>, timeout_microseconds@entry=30000000, working_dur=working_dur@entry=0x7f04da9085e0) at /usr/src/debug/ceph-14.0.1-3134-gdfall176/src/msg/async/Event.cc:401
#3 0x00007f04ec584815 in operator() (__closure=0x55c9d08e8758) at /usr/src/debug/ceph-14.0.1-3134-gdfall176/src/msg/async/Stack.cc:53
#4 std::_Function_handler<void(), NetworkStack::add_thread(unsigned int)::<lambda()>>::_M_invoke(const std::_Any_data &) (__functor=...) at /opt/rh/devtoolset-7/root/usr/include/c++/7/bits/std_function.h:316
#5 0x00007f04ec7cec5f in execute_native_thread_routine () from /usr/lib64/ceph/libceph-common.so.1
#6 0x00007f04e8c58dd5 in start_thread () from /lib64/libpthread.so.0
#7 0x00007f04e7b1eead in clone () from /lib64/libc.so.6
```

Thread 1 (Thread 0x7f04f4f75fc0 (LWP 151405)):

```
#0 0x00007f04e8c6049b in raise () from /lib64/libpthread.so.0
#1 0x000055c9ccc23da5 in reraise_fatal (signal=11) at /usr/src/debug/ceph-14.0.1-3134-gdfall176/src/global/signal_handler.cc:81
#2 handle_fatal_signal (signal=11) at /usr/src/debug/ceph-14.0.1-3134-gdfall176/src/global/signal_handler.cc:298
#3 <signal handler called>
#4 0x00007f04e8c59f47 in pthread_join () from /lib64/libpthread.so.0
#5 0x00007f04ec31261d in Thread::join (this=this@entry=0x55c9cfa14ea8, prval=prval@entry=0x0) at /usr/src/debug/ceph-14.0.1-3134-gdfall176/src/common/Thread.cc:168
#6 0x000055c9ccaef6a3 in RocksDBStore::close (this=0x55c9cfa14c80) at /usr/src/debug/ceph-14.0.1-3134-gdfall176/src/kv/RocksDBStore.cc:658
#7 0x000055c9ccaf0892 in RocksDBStore::~RocksDBStore (this=0x55c9cfa14c80, __in_chrg=<optimized out>) at /usr/src/debug/ceph-14.0.1-3134-gdfall176/src/kv/RocksDBStore.cc:628
#8 0x000055c9ccaf0c59 in RocksDBStore::~RocksDBStore (this=0x55c9cfa14c80, __in_chrg=<optimized out>) at /usr/src/debug/ceph-14.0.1-3134-gdfall176/src/kv/RocksDBStore.cc:648
#9 0x000055c9cc8b51fc in main (argc=<optimized out>, argv=<optimized out>) at /usr/src/debug/ceph-14.0.1-3134-gdfall176/src/ceph_mon.cc:846
```

#3 - 02/05/2019 07:15 PM - David Zafman

- Assignee set to David Zafman

#4 - 02/05/2019 11:27 PM - David Zafman

<https://github.com/ceph/ceph/pull/26290> has a workaround for this issue.

#5 - 02/06/2019 12:30 PM - Nathan Cutler

- Subject changed from *osd-backfill-space.sh exposes rocksdb hang* to *osd-backfill-space.sh exposes rocksdb hang*

#6 - 02/06/2019 04:44 PM - David Zafman

- Assignee deleted (David Zafman)

- Priority changed from Urgent to High

Workaround merged, so changed priority from Urgent to High.

#7 - 02/06/2019 10:06 PM - Neha Ojha

- Status changed from 12 to Resolved

#8 - 02/08/2019 03:56 PM - David Zafman

- Status changed from Resolved to 12

The workaround was for a specific case but I'm seeing lots more failures in master:

http://pulpito.ceph.com/dzafman-2019-02-07_22:34:38-rados-wip-dzafman-testing3-distro-basic-smithi/

Addendum:

Only job 3564441 is definitely this case

#9 - 02/09/2019 01:49 AM - David Zafman

Also, seen in http://qa-proxy.ceph.com/teuthology/dzafman-2019-02-07_22:34:38-rados-wip-dzafman-testing3-distro-basic-smithi/3564441

but in qa/standalone/osd/osd-recovery-prio.sh TEST_recovery_pool_priority() test

```
2019-02-08T09:00:01.920 INFO:tasks.workunit.client.0.smithi158.stderr:/home/ubuntu/ceph-test/clone.client.0/qa/standalone/osd/osd-recovery-prio.sh:492: TEST_recovery_pool_priority: kill_daemons td/osd-recovery-prio
2019-02-08T09:00:01.922 INFO:tasks.workunit.client.0.smithi158.stderr://home/ubuntu/ceph-test/clone.client.0/qa/standalone/ceph-helpers.sh:347: kill_daemons: shopt -q -o xtrace
2019-02-08T09:00:01.922 INFO:tasks.workunit.client.0.smithi158.stderr://home/ubuntu/ceph-test/clone.client.0/qa/standalone/ceph-helpers.sh:347: kill_daemons: echo true
2019-02-08T09:00:01.923 INFO:tasks.workunit.client.0.smithi158.stderr:/home/ubuntu/ceph-test/clone.client.0/qa/standalone/ceph-helpers.sh:347: kill_daemons: local trace=true
2019-02-08T09:00:01.923 INFO:tasks.workunit.client.0.smithi158.stderr:/home/ubuntu/ceph-test/clone.client.0/qa/standalone/ceph-helpers.sh:348: kill_daemons: true
2019-02-08T09:00:01.923 INFO:tasks.workunit.client.0.smithi158.stderr:/home/ubuntu/ceph-test/clone.client.0/qa/standalone/ceph-helpers.sh:348: kill_daemons: shopt -u -o xtrace
2019-02-08T09:06:05.260 INFO:tasks.workunit.client.0.smithi158.stderr:/home/ubuntu/ceph-test/clone.client.0/qa/standalone/ceph-helpers.sh:364: kill_daemons: return 1
2019-02-08T09:06:05.260 INFO:tasks.workunit.client.0.smithi158.stderr:/home/ubuntu/ceph-test/clone.client.0/qa/standalone/osd/osd-recovery-prio.sh:492: TEST_recovery_pool_priority: return 1
2019-02-08T09:06:05.260 INFO:tasks.workunit.client.0.smithi158.stderr:/home/ubuntu/ceph-test/clone.client.0/qa/standalone/osd/osd-recovery-prio.sh:37: run: return 1
2019-02-08T09:06:05.260 INFO:tasks.workunit.client.0.smithi158.stderr:/home/ubuntu/ceph-test/clone.client.0/qa/standalone/ceph-helpers.sh:2017: main: code=1
```

```
2019-02-08 09:00:02.860 7f15ec13f700 4 rocksdb: [/build/ceph-14.0.1-3501-g910a95b/src/rocksdb/db/compaction_j
ob.cc:1390] [default] [JOB 4] Compacted 2@0 files to L6 => 23818209 bytes
2019-02-08 09:00:02.860 7f15ec13f700 3 rocksdb: [/build/ceph-14.0.1-3501-g910a95b/src/rocksdb/db/version_set.
cc:2578] More existing levels in DB than needed. max_bytes_for_level_multiplier may not be guaranteed.
2019-02-08 09:00:02.920 7f15ec13f700 4 rocksdb: (Original Log Time 2019/02/08-09:00:02.921675) [/build/ceph-1
4.0.1-3501-g910a95b/src/rocksdb/db/compaction_job.cc:752] [default] compacted to: base level 6 max bytes base
268435456 files[0 0 0 0 0 1] max score 0.00, MB/sec: 58.7 rd, 58.6 wr, level 6, files in(2, 0) out(1) MB in(
22.7, 0.0) out(22.7), read-write-amplify(2.0) write-amplify(1.0) OK, records in: 2145, records dropped: 255 ou
tput_compression: NoCompression

2019-02-08 09:00:02.920 7f15ec13f700 4 rocksdb: (Original Log Time 2019/02/08-09:00:02.921710) EVENT_LOG_v1 {
"time_micros": 1549616402921696, "job": 4, "event": "compaction_finished", "compaction_time_micros": 406173, "
output_level": 6, "num_output_files": 1, "total_output_size": 23818209, "num_input_records": 2145, "num_output
_records": 1890, "num_subcompactions": 1, "output_compression": "NoCompression", "num_single_delete_mismatches
": 0, "num_single_delete_fallthrough": 0, "lsm_state": [0, 0, 0, 0, 0, 1]}
2019-02-08 09:00:02.924 7f15ec13f700 4 rocksdb: EVENT_LOG_v1 {"time_micros": 1549616402928073, "job": 4, "eve
nt": "table_file_deletion", "file_number": 10}
2019-02-08 09:00:02.924 7f15ec13f700 4 rocksdb: EVENT_LOG_v1 {"time_micros": 1549616402928159, "job": 4, "eve
nt": "table_file_deletion", "file_number": 4}
END OF LOG
```

#10 - 03/25/2019 03:04 PM - David Zafman

Seen in mimic backport testing with new osd-backfill-prio.sh test.

http://pulpito.ceph.com/dzafman-2019-03-20_19:47:51-rados-wip-zafman-testing-mimic-distro-basic-smithi/3754185

#11 - 05/08/2019 04:53 PM - Neha Ojha

another instance in mimic: /a/yuriw-2019-05-07_14:33:13-rados-wip-yuri-testing-2019-05-06-2158-mimic-distro-basic-smithi/3937276/

#12 - 08/21/2019 07:39 PM - David Zafman

- Priority changed from High to Normal

#13 - 12/05/2019 09:36 PM - Patrick Donnelly

- Status changed from 12 to New